

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**A Statistical Bigram Correlation Model
for Image Retrieval**

Inventors:

Mingjing Li

Zheng Chen

Liu Wenyin

Hong-Jiang Zhang

1 **TECHNICAL FIELD**

2 The following description relates to image or electronic image retrieval.

3
4 **BACKGROUND**

5 Digital images are increasingly more common as scanners and digital
6 cameras drop in price and increase in availability and function. As users such as
7 digital photographers, artists, and so on, amass large collections of digital
8 photographs on their computers, the challenges involved with querying and
9 accessing digital images on local and networked computing systems increase.
10 Thus, digital image users increasingly rely on conventional image retrieval
11 technology to help query and access digital images from various data stores. Such
12 image retrieval technology includes keyword-based image retrieval or content-
13 based image retrieval.

14 Keyword-based image retrieval finds images by matching keywords from a
15 user query to keywords that have been manually added to the images. Thus, these
16 images have been manually annotated with keywords related to their semantic
17 content. One of the more popular collections of annotated images is "Corel™
18 Gallery", an image database from Corel Corporation that includes upwards of one
19 million annotated images.

20 Unfortunately, with keyword-based image retrieval systems, it can be
21 difficult or impossible for a user to precisely describe the inherent complexity of
22 certain images. Additionally, image annotation is a subjective process—what may
23 be important to one user may not be important to another. As a result, retrieval
24 accuracy can be severely limited because some images—those that cannot be
25 described or can only be described ambiguously—will not be retrieved

1 successfully. In addition, due to the enormous burden of manual annotation, there
2 are a limited number of databases with annotated images.

3 Although image retrieval techniques based on keywords can be easily
4 automated, they suffer from the same problems as the information retrieval
5 systems in text databases and web-based search engines. Because of wide spread
6 synonymy and polysemy in natural language, the precision of such systems is very
7 low and their recall is inadequate. (Synonymy is the quality of being synonymous;
8 equivalence of meaning. Polysemy means having or characterized by many
9 meanings). In addition, linguistic barriers and the lack of uniform textual
10 descriptions for common image attributes severely limit the applicability of the
11 keyword based systems.

12 Content-based image retrieval (CBIR) systems have been built to address
13 many issues, such as those of keyword-based systems. These systems extract
14 visual image features such as color, texture, and shape from the image collections
15 and utilize them for retrieval purposes. These visual image features are also called
16 "low-level" features. Examples of low-level features of an image include color
17 histogram, wavelet based texture descriptors, directional histograms of edges, and
18 so forth. CBIR systems work well when the extracted feature vectors accurately
19 capture the essence of the image content.

20 For example, if a user is searching for an image with complex textures
21 having a particular combination of colors, this type of query is extremely difficult
22 to describe using keywords, but it can be reasonably represented by a combination
23 of color and texture features. On the other hand, if a user is searching for an
24 object that has clear semantic meanings but cannot be sufficiently represented by
25 combinations of available feature vectors, the content-based systems will not

1 return many relevant results. Furthermore, the inherent complexity of the images
2 makes it almost impossible for users to present the system with a query that fully
3 describes their intentions. Accordingly, although CBIR solves many of the
4 problems of keyword-based image retrieval, conventional CBIR technology has a
5 number of shortcomings.

6 One such shortcoming, for example, is that searches may return entirely
7 irrelevant images that just happen to possess similar features. Individual objects in
8 images contain a wide variety of low-level features. This increases the likelihood
9 that completely irrelevant images will be returned in response to a query that is
10 based on low-level features. Therefore, using only the low-level features of an
11 image to describe the types of images that the user wishes to locate will not
12 typically satisfactorily describe what a user desires to retrieve.

13 Another shortcoming, for example, is that users typically desire to locate
14 images that are based on specific semantic concepts, rather than images that
15 include certain low-level features. Semantic concepts include meaningful content
16 of an image—for example, a river, a person, a car, a boat, etc. Although
17 objectively measurable, low-level image features lack specific meaning.
18 Additionally, mapping semantic concepts to low-level features is still impractical
19 with present computer vision and AI techniques. Accordingly, the disparity
20 between semantic content and low-level features that lack specific meaning
21 substantially limits the performance of conventional CBIR systems.

22 To improve this situation, some CBIR systems utilize user feedback to gain
23 an understanding as to the relevancy of certain images. The user feedback is in the
24 form of selected exemplary images. These exemplary images may be called
25 “feedback” images. A user feedback selects such exemplary images to narrow

1007441-0010
2015-07-20 14:54:00

1 successive searches. A common approach to relevance feedback is estimating
2 ideal query parameters using the low-level image features of the exemplary
3 images. Thus, relevance feedback assists in mapping low-level features to human
4 recognition of semantic concepts.

5 In a relevance-feedback CBIR system, a user submits a query and the
6 system provides a set of query results. More specifically, after a query, the system
7 presents a set of images to the user. The user designates specific images as
8 positive or negative. Positive indicates that the image contains the semantic
9 concepts queried and negative indicates that the image does not contain such
10 concepts. Based upon this feedback, the system performs a new query and
11 displays a new set of resulting images. This means that relevance feedback is
12 dynamically used during the particular single search session to modify a search
13 query vector or distance metric, or to update a probability distribution of images
14 across a database.

15 Each round of query and feedback in a particular search session may be
16 called an iteration of that particular search session. This query/feedback process
17 continues for some number of iterations or until the user is either satisfied with the
18 overall relevance of the present set of images, or decides to attempt a different
19 search query. In this manner, image relevance feedback from the user may reveal
20 semantic relationships between the retrieved images that are not easily captured by
21 image low-level features.

22 Unfortunately, image relevance feedback is not typically accumulated or
23 memorized across CBIR search sessions. Rather, such image relevance feedback
24 is typically discarded and not utilized to improve future performance of the CBIR

1 system. The following arrangements and procedures address these and other
2 limitations of conventional CBIR techniques.

3 4 **SUMMARY**

5 The described arrangements and procedures for improving iterative results
6 of content-based image retrieval (CBIR) using a bigram model to correlate
7 relevance feedback. Specifically, multiple images are received responsive to
8 multiple image search sessions. Relevance feedback is used to determine whether
9 the received images are semantically relevant. A respective semantic correlation
10 between each of at least one pair of the images is then estimated using respective
11 bigram frequencies. The bigram frequencies are based on multiple search sessions
12 in which each image of a pair of images is semantically relevant.

13 14 **BRIEF DESCRIPTION OF THE DRAWINGS**

15 Fig. 1 illustrates an exemplary system that uses a bigram correlation model
16 to accumulate semantic relationships between images from image relevance
17 feedback information.

18 Fig. 2 shows an exemplary host computer that uses a bigram correlation
19 model to accumulate semantic relationships between images from user-provided
20 relevance feedback information.

21 Fig. 3 shows an exemplary procedure to utilize a bigram correlation model
22 to accumulate semantic relationships between images from user-provided
23 relevance feedback information.

1 Fig. 4 shows further features of an exemplary procedure to utilize a bigram
2 correlation model to accumulate semantic relationships between images from user-
3 provided relevance feedback information.

4 Fig. 5 shows an example of a suitable computing environment on which an
5 exemplary system and procedure to utilize a bigram correlation model to
6 accumulate semantic relationships between images from user-provided relevance
7 feedback information.

8 The same numbers are used throughout the drawings to reference like
9 features and components.

11 **DETAILED DESCRIPTION**

12 The following description sets forth exemplary subject matter to retrieve
13 semantically related images responsive to a search query. The subject matter is
14 described with specificity to meet statutory requirements. However, the
15 description itself is not intended to limit the scope of this patent. Rather, the
16 inventors have contemplated that the claimed subject matter might also be
17 embodied in other ways, to include different elements or combinations of elements
18 similar to the ones described in this document, in conjunction with other present or
19 future technologies.

20 **Incorporation by Reference**

21 The following co-pending patent applications assigned to the assignee
22 hereof, the Microsoft Corporation, are incorporated herein by reference:

- 23 • U.S. Patent Application Serial No. 09/702,292, entitled "Image
24 Retrieval Systems and Methods with Semantic and Feature Based
25 Relevance Feedback", filed on October 30, 2000;

- U.S. Patent Application Serial No. 09/702,288, entitled “Semi-Automatic Annotation of Multimedia Objects”, filed on October 30, 2000; and
- U.S. Patent Application Serial No. 09/823,534, entitled “Relevance Maximizing, Iteration Minimizing, Relevance-Feedback, Content-Based Image Retrieval (CBIR)”, filed on March 30, 2001.

Overview

A statistical bigram correlation model for image retrieval is disclosed to accumulate semantic relationships between images from user-provided relevance feedback information. This accumulated information is incorporated into an image retrieval system so that it can be used across multiple search sessions to retrieve semantically consistent images. Specifically, responsive to obtaining results of a search session, probabilities are determined indicating whether images are semantically similar to one another based on the co-occurrence frequency that the images were identified as relevant images during a previous query/feedback session. Such probabilities are dynamically updated in the system during the searching process and can also be trained from user relevance feedback logs.

An Exemplary System

Fig. 1 illustrates an exemplary system that uses a bigram correlation model to accumulate semantic relationships between images based on user-provided relevance feedback information. In environment 100 one or more (x) clients 102 are coupled to a media content store 104. The media content store 104 is any combination of local storage (e.g., local volatile or non-volatile memory), networked storage (e.g., a parallel connection, an organizational intranet network, the Internet, and so on), or other communication configurations.

1 These communication configurations provide for electronic exchange of
2 information using an appropriate protocol (e.g., TCP/IP, UDP, SOAP, etc.)
3 between the host device 102 and one or more media content sources or servers that
4 include multiple (y) pieces of media content 106. This electronic exchange
5 provides for client 102 communication with media content store 104 to access
6 (e.g., view, search, download, etc.) pieces of media content 106.

7 The storage of media content pieces 106 within media content store 104 can
8 be arranged in any of a wide variety of manners and according to any of a wide
9 variety of data formats. For example, media content pieces 106 may be stored on
10 multiple servers hosting Web pages accessible via a network using an appropriate
11 protocol such as Hypertext Transfer Protocol (HTTP). Web pages are documents
12 that a user can view or otherwise render and which typically include links to one
13 or more other pages that the user can access. Web pages are typically stored as
14 one or more files at a remote location(s), being accessed by the user via a
15 computer that is operatively coupled to a network. Web pages often include
16 multiple pieces of media content 106.

17 Media content pieces 106 include still images, frames of motion video,
18 audio, multimedia, and so on. A piece of media content 106 refers to media
19 content that can be rendered such as a single visual image, and the like.

20 A user of a client 102 searches the media content store 104 for pieces of
21 media content 106. As a user operates within the computing environment of a
22 client 102, the client 102 monitors the user's search session activities and detects a
23 user's relevance feedback that indicates which of a number of pieces of media
24 content 106 are relevant to a search session. The client 102 uses a statistical
25 bigram correlation model to accumulate semantic relationships between images

1 from user-provided relevance feedback information. Aspects of the bigram
2 correlation model are described in greater detail below in reference to Fig. 2. This
3 accumulated information can be used across multiple image search sessions
4 conducted on the client 102 to retrieve semantically consistent images
5 corresponding to a respective search operation.

6 **A Client Computer 102**

7 Fig. 2 shows an exemplary computing device 102 that uses a bigram
8 correlation mode to accumulate semantic relationships between images from user-
9 provided relevance feedback information. The computer 102 is operational as any
10 one of a number of different computing devices such as a personal computer, an
11 image server computer, a thin client, a thick client, a hand-held or laptop device, a
12 multiprocessor system, a microprocessor-based system, a set top box,
13 programmable consumer electronics, a wireless phone, an application specific
14 integrated circuit (ASIC), a network PC, minicomputer, mainframe computer, and
15 so on.

16 The host computer includes a processor 202 that is coupled to a system
17 memory 204. The system memory 204 includes any combination of volatile and
18 non-volatile computer-readable media for reading and writing. Volatile computer-
19 readable media includes, for example, random access memory (RAM). Non-
20 volatile computer-readable media includes, for example, read only memory
21 (ROM), magnetic media such as a hard-disk, an optical disk drive, a floppy
22 diskette, a flash memory card, a CD-ROM, and so on.

23 The processor 202 is configured to fetch and execute computer program
24 instructions from program modules 206; and configured to fetch data 208 while
25 executing the program modules 206. Program modules typically include routines,

1 programs, objects, components, data structures, etc., for performing particular
2 tasks or implementing particular abstract data types.

3 Program modules 206 include the search engine module 210, a user
4 relevance feedback module 212, a correlation analysis module 214, an off-line
5 training module 216, an on-line training module 218, a Web browser module 220,
6 an operating system (not shown), and so on. Program data 208 includes, image
7 data 222, correlation data 224 (e.g., a bigram frequency, a unigram frequency, a
8 maximum frequency, a self correlation value, a mutual correlation value, and so
9 on), and other data 226 (e.g., a number of search sessions conducted on the
10 client 102, user log of image relevance feedback, etc.), and so on.

11 The search engine 210 search session starts with a query phase, which is
12 possibly followed by one or more user feedback and image correlation phases.
13 The search engine 210 or query handler receives a search query that is generated
14 from information input by a user. Such queries may be in the form of natural
15 language queries, individual word queries, or image queries that contains low-
16 level features of an example image that forms the basis of the search.

17 Natural language queries and individual word queries include a textual
18 description of the search criteria pertaining to the types of images desired. Textual
19 description is converted to a text feature vector by extracting keywords and stored
20 as a query vector 226. If an image is used for the search criteria, low level
21 features of the image are extracted and used to generate the initial query feature
22 vector 226. Depending on the query type, the query handler 210 initiates either a
23 keyword-based or feature-based search of the media content store 104 of Fig. 1.

24 The user relevance feedback module 212 displays at least a portion of the
25 images 222 returned by the search engine 210 for user feedback. The feedback

1 module 212 ranks the retrieved images according to their relevance to the
2 query 226. If no previous user feedback 228 has yet been acquired, the
3 images 222 are ranked based on the similarity of the images to feature vectors in
4 the query. As discussed in greater detail below in reference to the correlation
5 module 214, the final ranking score for each retrieved image 222 image is the
6 weighted sum of the feature similarity measure (i.e., with respect to the search
7 query) and semantic support for the image. A display device 230 operatively
8 coupled to the computer 102 displays the ranked images 222 via a user interface
9 with which the user can mark or otherwise identify individual images as more,
10 less, or not relevant to the query. The user feedback module 212 records such user
11 feedback into a user log 228, which indicates those images deemed relevant to one
12 or more search sessions (i.e., positive feedback) and which images are not (i.e.,
13 negative feedback). Only those images with a highest range of ranking scores may
14 be displayed. Such a range can be determined with a predefined threshold or by
15 determining that only a fixed number of images will be retrieved.

16 The correlation module 214 imposes semantic constraints to the
17 image(s) 222 retrieved by the search engine 210 in response to the user input
18 search query 226. Any imposed semantic constraints are based on previously
19 provided (i.e., identified in a user log 228) or presently provided user image
20 relevance feedback information. That is, each retrieved image 222 is assigned a
21 semantic support (i.e., the semantic correlation value 224).

22 Initially, the semantic support of an image is set to its feature-based
23 similarity:

24
$$P(I) = S(I), \text{ where } 0 \leq S(I) \leq 1 \text{ is the similarity of Image } I.$$

1 If the user provides any relevance feedback via the feedback module 212,
2 the similarity measure $S(I)$ is refined accordingly, and the images are re-ranked.
3 For instance, the similarity measure of relevant images is set to 1, while that of
4 irrelevant ones is set to 0, and that of other images is recalculated. In this way,
5 contribution from non-feedback ones in the retrieved list of images is discounted
6 because of their uncertainty in terms of semantic similarities to the query. Then
7 the semantic support $P(I)$ is iteratively updated through the correlation model for a
8 number of k times according to the following formulas:

$$9 \quad P'(I) = \frac{\sum_{j=1}^M P(I_j) * R(I, I_j)}{\sum_{j=1}^M P(I_j)},$$

$$10 \quad P(I) = P'(I),$$

11 where $R(I, I_j)$ is the correlation between image I and I_j , $I_j (j=1, \dots, M)$ are M
12 images with the highest similarities. (Various formulas for determining $R(I, I_j)$
13 are discussed below). The final ranking score (i.e., other data 228) of each
14 retrieved image 222 is then the weighted sum of the calculated feature similarity
15 measure and the semantic support:

$$16 \quad Score(I) = w * P(I) + (1 - w) * S(I), \quad 0 \leq w \leq 1,$$

17 where $S(I)$ is the similarity measure of image I , $P(I)$ is its semantic support, w is
18 the semantic weight. Images with the highest scores are returned to the user as the
19 final retrieval results.

20 Semantic Correlation between Images

21 The correlation module 214 estimates the semantic correlation between two
22 images 222 based on the number of search sessions in which both images are
23 marked by the user (via the feedback module 212) as being relevant to the search

1 session. The number of search sessions in which two images are jointly labeled as
2 relevant is referred to as *bigram frequency* 224. The number of search sessions
3 wherein an image is labeled as relevant is referred to as *unigram frequency* 224.

4 The maximum value of all unigram and bigram frequencies is referred to as
5 maximum frequency 224. The mutual correlation 224, i.e., the correlation
6 between two different images, is defined as the ratio between their bigram
7 frequency and the maximum frequency 224. The self-correlation 224, i.e., the
8 correlation between an image 222 and itself, is defined as the ratio between its
9 unigram frequency 224 and the maximum frequency 224. Since the bigram
10 frequency 224 is symmetric, the defined semantic correlation 224 is also
11 symmetric. Thus, a triangular matrix is used to represent the correlation model of
12 correlation module 214.

13 To fully utilize the information provided by the relevance feedback
14 module 212, and to reflect the diversity of users' search intentions, the respective
15 definitions of unigram and bigram frequencies 224 are extended to take account of
16 irrelevant images. Specifically, there is a positive correlation between two
17 relevant images, and a negative correlation between a relevant image and an
18 irrelevant image, but no correlation otherwise. In case that the value of a bigram
19 or unigram frequency is less than zero, the corresponding correlation value is set
20 to zero.

21 For instance, the semantic correlation R between two images I and J can be
22 determined as follows:

- 23 • $0 \leq R(I, J) \leq 1$ (attributes);
- 24 • $R(I, J) = R(J, I)$ (attributes);
- 25 • if $I=J$ and $U(I) \leq 0$: $R(I, J) = 0$ (attributes);

- if $I \neq J$ and $B(I, J) \leq 0$: $R(I, J) = 0$ (attributes);
- if $I = J$ and $U(I) > 0$: $R(I, J) = U(I)/T$ (self correlation); or
- if $I \neq J$ and $B(I, J) > 0$: $R(I, J) = B(I, J)/T$ (mutual correlation).

where I, J are two images, $B(I, J)$ is their bigram frequency, $U(I)$ is the unigram frequency of image I , T is the maximum frequency, $R(I, J)$ is the correlation between image I and J .

The correlation module 212 stores calculated semantic correlation data 224 into the system such as in a user log 228.

Offline Training

The offline training module 216 calculates the unigram and bigram frequencies 224 from the relevance feedback information collected in a user log (i.e., stored in other data 228). Initially, all unigram and bigram frequencies 224 are set to equal zero (0). To overcome data sparseness, search sessions with the same query, either a text query or an image example, are grouped together such that feedback images 220 in different search sessions may obtain correlation information. Within each group of search sessions with the same query, the unigram counts 224 are calculated. Based on these counts 224, the unigram and bigram frequencies 224 are updated accordingly.

The unigram count 224 in a group is calculated as follows. At first, $C(I)$ is set to 0, where $C(I)$ is the unigram count of image I . After that, $C(I)$ is iteratively updated for every session in this group: $C(I) = C(I) + 1$, if image I is labeled as relevant in a session; $C(I) = C(I) - 1$, if image I is labeled as irrelevant in a session; $C(I)$ is unchanged otherwise. This process is repeated for every image in the database 222.

1 The unigram frequencies 224 are updated as: $U(I) = U(I) + C(I)$. The
2 bigram frequencies 224 of image pairs are updated as:

- 3 • $B(I, J) = B(I, J) + \min\{C(I), C(J)\}$, if $C(I) > 0, C(J) > 0$,
- 4 • $B(I, J) = B(I, J) - \min\{C(I), -C(J)\}$, if $C(I) > 0, C(J) < 0$,
- 5 • $B(I, J) = B(I, J) - \min\{-C(I), C(J)\}$, if $C(I) < 0, C(J) > 0$, or
- 6 • $B(I, J) = B(I, J)$, otherwise.

7 Finally, the correlation value 224 is determined as discussed above.

8 On-Line Training

9 The online training module 218 dynamically updates the unigram and
10 bigram frequencies 224 with relevance feedback information (i.e., via the
11 relevance feedback module 212) in the current search session (i.e., see, the search
12 engine 210) at the end of each session. Initially, all unigram and bigram
13 frequencies 224 are calculated by offline training module 216 from the user
14 log 226. In this manner, user relevance feedback 224 is utilized across multiple
15 image search sessions. These frequencies 224 are set to 0 if there is no log 226.
16 This algorithm is similar to that of offline training, except that a session group
17 only contains one search session.

18 The unigram count 224 is calculated as:

- 19 • $C(I) = 1$ if I is relevant;
- 20 • $C(I) = -1$ if I is irrelevant; and
- 21 • $C(I) = 0$ if I is a non-feedback image.

22 The unigram frequencies 224 are updated as: $U(I) = U(I) + C(I)$.

23 The bigram frequencies 224 of image pairs are updated as:

- 24 • $B(I, J) = B(I, J) + 1$, if $C(I) > 0, C(J) > 0$,

- $B(I, J) = B(I, J) - 1$, if $C(I) > 0, C(J) < 0$,
- $B(I, J) = B(I, J) - 1$, if $C(I) < 0, C(J) > 0$, or
- $B(I, J) = B(I, J)$, otherwise.

After that, any affected semantic correlations 224 are updated as discussed above.

An Exemplary Procedure to Retrieve Images for One Search Session

Initially, all unigram and bigram frequencies are set to zero (0). At block 302, the image correlation model is optionally trained off-line. To overcome a sparse data set, search sessions (block 304) with a same query 226 (i.e., either a text query or an image example query) are grouped together such that feedback images (i.e., relevant, non-relevant, and/or unchanged images) in different sessions may obtain semantic correlation information saved from previous search sessions. Within each group of search sessions with the same query, unigram counts 224 are calculated. Based on these counts 224, the unigram and bigram frequencies 224 are updated accordingly.

Specifically, the off-line correlation training operations are as follows:

- (a) initialize all unigram and bigram frequencies to zero;
- (b) cluster search sessions with a same query into groups;
- (c) calculate the unigram counts within a group;
- (d) update the unigram frequencies ;
- (e) update the bigram frequencies;
- (f) repeat operations (c), (d), and (e) for all session groups;
- (g) set all negative unigram and bigram frequencies to zero; and
- (h) calculate the correlation values $R(I, J)$ according to the above described formulas.

1 The offline training of the correlation model is optional, because if there is
2 no previous user feedback or user log, the bigram and unigram frequencies are set
3 to zero during online training.

4 At block 304, the procedure 300 performs feature-based image search and
5 retrieval based on a search query 226, which can be either text or image based, or
6 based on provided user feedback (block 402 of Fig. 4). At block 306 the
7 procedure 300 determines whether the user has provided image relevance
8 feedback (i.e., prior user feedback is stored in a user log 228) during/after prior
9 search sessions 304. If the user has not provided previous image relevance
10 feedback, at block 308, the procedure 300 displays ranked images for user
11 feedback and/or selection. The procedure 300 continues at online reference "B" of
12 Fig. 4.

13 At block 310, it has been determined that they user has provided previous
14 image relevance feedback (block 306), the procedure 300 provides semantic
15 support for each of the retrieved images (block 304) based on the values in the
16 user log 228. At block 312, the procedure 300 re-ranks, reorders, or scores
17 ($Score(I)$) the images based on image similarity measure $S(I)$, semantic support
18 $P(I)$, and semantic weight w . At block 308 the procedure 300 displays the ranked
19 images for user feedback. The procedure 300 continues at online reference "B" of
20 Fig. 4.

21 Fig. 4 shows further features of an exemplary procedure 300 to utilize a
22 bigram correlation of relevance feedback for image retrieval. At block 402, the
23 procedure 300 determines whether the user has provided additional relevance
24 feedback with respect to the ranked images presently being displayed (block 306
25 of Fig. 3). If so, the procedure 300 continues at page reference "A" of Fig. 3.

1 At block 404, the procedure 300 updates the cached unigram and bigram
2 frequencies and the correlation model based on the feedback of the current
3 session.

4 **Exemplary Computing Environment**

5 Fig. 5 shows an example of a suitable computing environment on which an
6 exemplary system and procedure to utilize a bigram correlation of relevance
7 feedback for image retrieval may be implemented. Exemplary computing
8 environment 500 is only one example of a suitable computing environment and is
9 not intended to suggest any limitation as to the scope of use or functionality of an
10 exemplary system and procedure to cluster queries. The computing environment
11 500 should not be interpreted as having any dependency or requirement relating to
12 any one or combination of components illustrated in the exemplary computing
13 environment 500.

14 An exemplary system and procedure to improve iterative results of CBIR
15 using a bigram model to correlate relevance feedback may be described in the
16 general context of computer-executable instructions, such as program modules,
17 being executed by a computer. Generally, program modules include routines,
18 programs, objects, components, data structures, etc., that perform particular tasks
19 or implement particular abstract data types. An exemplary system and procedure
20 to improve iterative results of CBIR using a bigram model to correlate relevance
21 feedback may also be practiced in distributed computing environments where
22 tasks are performed by remote processing devices that are linked through a
23 communications network. In a distributed computing environment, program
24 modules may be located in both local and remote computer storage media
25 including memory storage devices.

1 As shown in Fig. 5, the computing environment 500 includes a
2 general-purpose computing device in the form of a computer 102 of Figs. 1 and 2.
3 The components of computer 102 may include, by are not limited to, one or more
4 processors or processing units 202, a system memory 204, and a bus 516 that
5 couples various system components including the system memory 204 to the
6 processor 202.

7 Bus 516 represents one or more of any of several types of bus structures,
8 including a memory bus or memory controller, a peripheral bus, an accelerated
9 graphics port, and a processor or local bus using any of a variety of bus
10 architectures. By way of example, and not limitation, such architectures include
11 Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA)
12 bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA)
13 local bus, and Peripheral Component Interconnects (PCI) bus also known as
14 Mezzanine bus.

15 Computer 102 typically includes a variety of computer-readable media.
16 Such media may be any available media that is accessible by the computer 102,
17 and it includes both volatile and non-volatile media, removable and non-
18 removable media. For example, the system memory 204 includes computer
19 readable media in the form of volatile memory, such as random access memory
20 (RAM) 520, and/or non-volatile memory, such as read only memory (ROM) 518.
21 A basic input/output system (BIOS) 522, containing the basic routines that help to
22 transfer information between elements within computer 102, such as during start-
23 up, is stored in ROM 518. RAM 520 typically contains data 208 and/or program
24 modules 206 that are immediately accessible to and/or presently be operated on by
25 processor 202.

Computer 102 may further include other removable/non-removable, volatile/non-volatile computer storage media. By way of example only, Fig. 5 illustrates a hard disk drive 524 for reading from and writing to a non-removable, non-volatile magnetic media (not shown and typically called a "hard drive"), a magnetic disk drive 526 for reading from and writing to a removable, non-volatile magnetic disk 528 (e.g., a "floppy disk"), and an optical disk drive 530 for reading from or writing to a removable, non-volatile optical disk 532 such as a CD-ROM, DVD-ROM or other optical media. The hard disk drive 524, magnetic disk drive 526, and optical disk drive 530 are each connected to bus 516 by one or more interfaces 534.

The drives and their associated computer-readable media provide nonvolatile storage of computer readable instructions, data structures, program modules, and other data for computer 102. Although the exemplary environment described herein employs a hard disk, a removable magnetic disk 528 and a removable optical disk 532, it should be appreciated by those skilled in the art that other types of computer readable media which can store data that is accessible by a computer, such as magnetic cassettes, flash memory cards, digital video disks, random access memories (RAMs), read only memories (ROM), and the like, may also be used in the exemplary operating environment.

A number of program modules may be stored on the hard disk, magnetic disk 528, optical disk 532, ROM 518, or RAM 520, including, by way of example, and not limitation, an OS 538, one or more application programs 206, other program modules 542, and program data 208. Each such OS 538, one or more application programs 206, other program modules 542, and program data 208 (or some combination thereof) may include an embodiment of an exemplary system

1 and procedure to improve iterative results of CBIR using a bigram model to
2 correlate relevance feedback.

3 A user may enter commands and information into computer 102 through
4 input devices such as keyboard 546 and pointing device 548 (such as a "mouse").
5 Other input devices (not shown) may include a microphone, joystick, game pad,
6 satellite dish, serial port, scanner, or the like. These and other input devices are
7 connected to the processing unit 202 through a user input interface 550 that is
8 coupled to bus 516, but may be connected by other interface and bus structures,
9 such as a parallel port, game port, or a universal serial bus (USB).

10 A monitor 552 (e.g., the monitor 230 of Fig. 2) or other type of display
11 device is also connected to bus 516 via an interface, such as a video adapter 554.
12 In addition to the monitor, personal computers typically include other peripheral
13 output devices (not shown), such as speakers and printers, which may be
14 connected through output peripheral interface 555.

15 Computer 102 may operate in a networked environment using logical
16 connections to one or more remote computers, such as a remote computer 562.
17 Logical connections shown in Fig. 5 are a local area network (LAN) 557 and a
18 general wide area network (WAN) 559. Such networking environments are
19 commonplace in offices, enterprise-wide computer networks, intranets, and the
20 Internet. Remote computer 562 may include many or all of the elements and
21 features described herein relative to computer 102.

22 When used in a LAN networking environment, the computer 102 is
23 connected to LAN 557 via network interface or adapter 566. When used in a
24 WAN networking environment, the computer typically includes a modem 558 or
25 other means for establishing communications over the WAN 559. The

modem 558, which may be internal or external, may be connected to the system bus 516 via the user input interface 550 or other appropriate mechanism.

Depicted in Fig. 5 is a specific implementation of a WAN via the Internet. Computer 102 typically includes a modem 558 or other means for establishing communications over the Internet 560. Modem 558, which may be internal or external, is connected to bus 516 via interface 550.

In a networked environment, program modules depicted relative to the personal computer 102, or portions thereof, may be stored in a remote memory storage device. By way of example, and not limitation, Fig. 5 illustrates remote application programs 569 as residing on a memory device of remote computer 562. The network connections shown and described are exemplary and other means of establishing a communications link between the computers may be used.

Computer Readable Media

An implementation of exemplary subject matter to system and procedure to improve iterative results of CBIR using a bigram model to correlate relevance feedback may be stored on or transmitted across some form of computer-readable media. Computer-readable media can be any available media that can be accessed by a computer. By way of example, and not limitation, computer readable media may comprise "computer storage media" and "communications media."

"Computer storage media" include volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules, or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic

1 tape, magnetic disk storage or other magnetic storage devices, or any other
2 medium which can be used to store the desired information and which can be
3 accessed by a computer.

4 “Communication media” typically embodies computer readable
5 instructions, data structures, program modules, or other data in a modulated data
6 signal, such as carrier wave or other transport mechanism. Communication media
7 also includes any information delivery media.

8 The term “modulated data signal” means a signal that has one or more of its
9 characteristics set or changed in such a manner as to encode information in the
10 signal. By way of example, and not limitation, communication media includes
11 wired media such as a wired network or direct-wired connection, and wireless
12 media such as acoustic, RF, infrared, and other wireless media. Combinations of
13 any of the above are also included within the scope of computer readable media.

14 **Conclusion**

15 The described arrangements and procedures provide for a bigram
16 correlation of relevance feedback for image retrieval. Although the arrangements
17 and systems to improve iterative results of CBIR using a bigram model to
18 correlate relevance feedback have been described in language specific to structural
19 features and methodological operations, it is to be understood that the
20 arrangements and procedures as defined the appended claims are not necessarily
21 limited to the specific features or operations described. Rather, the specific
22 features and operations are disclosed as preferred forms of implementing the
23 claimed subject matter.